Der ***t*-Test** bezeichnet eine Gruppe von statistischen Test. Zur Durchführbarkeit müssen folgende Voraussetzungen erfüllt sein:

- Die Varianzen s^2 der beiden Stichproben unterscheiden sich nicht *signifikant.* Dies lässt sich mit dem **F-Test** prüfen. Eine Abwandlung des t-Tests, der etwas weniger bekannte WELCH-Test, setzt keine gleiche Varianz voraus.
- Der Stichprobenumfang ist genügend groß um die Grundgesamtheit zu repräsentieren.
- Die Stichprobenwerte entstammen aus einer normalverteilten Grundgesamtheit.

Es werden folgende t-Tests unterschieden:

- Der **Einstichproben-t-Test (= Einfacher t-Test)** prüft ob sich der Mittelwert einer Stichprobe von einem Sollwert unterscheidet. ***Beispiel einer Fragestellung:*** Die Packungen eines Paracetamol-Präparates besitzen laut Aufdruck 400 mg Paracetamol pro Tablette. Die stichprobenartige chemische Analyse ergab im Mittel jedoch 395 mg pro Tablette. Ist das nur Zufall und auf eine ungenügend hoher Zahl an Stichproben zurückzuführen oder ist tatsächlich weniger Paracetamol enthalten, als angegeben?
- Der **Zweistichproben-t-Test (auch Doppelter t-Test)** prüft anhand der Mittelwerte zweier Stichproben, wie sich die Mittelwerte zweier Grundgesamtheiten zueinander verhalten. ***Beispiel einer Fragestellung:*** 2 Packungen Paracetamol werden jeweils für sich stichprobenartig auf den Paracetamolgehalt geprüft und die Mittelwerte verglichen:

  . Packung A: 395 mg/Tablette. Packung B: 399 mg/Tablette

  Ist das nur Zufall und auf eine ungenügend hoher Zahl an Stichproben zurückzuführen, oder unterscheiden sich tatsächlich die wahren Paracetamolgehalte?
- Der **Abhängige t-Test (auch Paardifferenzentest)** prüft für zwei verbundene (abhängige) Stichproben, ob sich die mittlere Differenz der Messwerte sich von Wert Δ = 0 unterscheidet. Dabei wird vorausgesetzt, dass die Differenzen normalverteilt sind. ***Beispiel einer Fragestellung:*** 10 Personen mit Fieber wurde ein Paracetamolpräparat gegeben. Dabei sank das Fieber im Mittel um 1,3 °C. Ist das nur Zufall und auf eine ungenügend hoher Zahl an Stichproben zurückzuführen, oder besitzt das Paracetamol fiebersenkende Wirkung?

Die t-Tests laufen darauf hinaus, dass zuerst eine Behauptung (Hypothese) formuliert wird. Dies ist in der Regel die **Nullhypothese (H_0).:** ***„Wir geht davon aus, die sich die beiden Vergleichsgrößen nur zufällig unterscheiden und auf die Stichprobenbildung zurückzuführen sind“.***

Anschließend wird dann mit einem Rechenverfahren eine **Prüfgröße (= *Prüfstatistik = Teststatistik***) berechnet und ggf. mit einem tabellierten Wert verglichen. Überschreitet die berechnete Teststatistik (Prüfgröße) den tabellierten Wert, wird die Nullhypothese verworfen, d.h. man geht von einem tatsächlichen Unterschied aus. Diese Verwerfung unterliegt aber einer gewissen Irrtumswahrscheindlichkeit (α). Endgültige Gewissheit, ob sich die tatsächlichen Werte unterscheiden, bekommt man mit dem t-Test also nicht. Je nachdem welche Irrtumswahrscheinlichkeit (α) man sich einräumt, muss man seine Teststatistik mit einem anderen Tabellenwert vergleichen.

Räumt man sich eine Irrtumswahrscheinlichkeit von unter α = 5% bei der Ablehnung der Nullhypothese, so sagt man auch: „Die Werte unterscheiden sich statistisch *signifikant*“. Ist man mit sich strenger und räumt sich nur eine Irrtumswahrscheinlichkeit von unter α = 1% ein, so sagt man, die Werte unterscheiden sich „statistisch *hoch signifikant*“

Tabellenkalkulationsprogramme können mit der Funktion TTEST (bzw. T.TEST) direkt die Irrtumswahrscheinlichkeit α berechnen.

t-Test zweier unabhängiger Stichprobenmittelwerte	**FTCCT_3_4**

Beispiel: Bei zwei Farbstoffproben wurden mehrmals der Gehalt bestimmt. Mit den Stichprobenumfängen n_A = 10 und n_B = 12 resultieren folgende Ergebnisse (in mg/L):

													Mittelwert	**Standardabweichung**
Probe A:	20,8	21,5	20,4	21,3	20,1	20,4	21,0	20,9	21,2	19,8			**20,74**	**0,550151494287407**
Probe B:	21,2	20,1	19,9	20,5	19,9	21,0	19,8	20,2	20,0	20,8	20,1	20,1	**20,30**	**0,465149047471492**

Es ist auf dem Signifikanzniveau von 95% (Irrtumswahrscheinlichkeit α = 5% = 0,05) zu prüfen, ob es einen tatsächlichen Unterschied zwischen den wahren Gehalten gibt.

Allgemeine Vorgehensweise	**Konkretisierung für dieses Beispiel**
1. Formulierung der Nullhypothese (H_0) Die Nullhypothese ist immer die Annahme, dass der Unterschied zwischen den Mittelwerten zufälliger Natur ist, die beiden *wahren Mittelwerte* (würde man erhalten wenn man die Mittelwerte aus sehr großen oder unendlich großen Stichprobenumfängen vergleichen würde) sind identisch.	**1. Formulierung der Nullhypothese (H_0)** Wir gehen davon aus, dass die mittlere Farbstoff-konzentration beider Proben in Wirklichkeit identisch ist. Der Unterschied zwischen den beiden Mittelwerten ist nur zufälliger Natur, da nur Stichproben miteinander ver-glichen wurden.
2. Berechnung der Teststatistik, $t_{Prüf}$: Die Berechnung erfolgt mithilfe der gemeinsamen Standardabweichung $s_{A,B}$: $$s_{A,B}=\pm\sqrt{\frac{(n_A-1)\cdot s_A^2+(n_B-1)\cdot s_B^2}{n_A+n_B-2}}$$ ODER wenn alle Messergebnisse x bekannt sind: $$s_{A,B}=\pm\sqrt{\frac{\sum(x_A-\bar{x}_A)^2+\sum(x_B-\bar{x}_B)^2}{n_A+n_B-2}}$$ und folgender Formel für die Teststatistik $$t_{Prüf}=\frac{\lvert\bar{x}_A-\bar{x}_B\rvert}{s_{A,B}}\cdot\sqrt{\frac{n_A\cdot n_B}{n_A+n_B}}$$	**2. Berechnung der Teststatistik, $t_{Prüf}$:** Da die Standardabweichungen s_A und s_B gegeben sind: $$t_{Prüf}\approx\frac{\lvert 20{,}74-20{,}30\rvert}{0{,}50517}\cdot\sqrt{\frac{10\cdot 12}{10+12}}\approx 2{,}0342$$
3. Vergleich mit Tabellenwert Der **Freiheitsgrad** beim t-Test zweier Stichproben ist $$f=n_1+n_2-2$$ Wenn für das gegebene **Signifikanzniveau** gilt: $t_{Prüf} > t_{Tabelle}$: Nullhypothese wird verworfen. $t_{Prüf} < t_{Tabelle}$: Nullhypothese wird akzeptiert.	**3. Vergleich mit Tabellenwert** Der Freiheitsgrad beträgt $$f=10+12-2=20$$ In der Tabelle findet sich in der Zeile f = 20 mit α = 0,05 der Wert $t_{Tabelle}$ = 2,086. Da $t_{Prüf} < t_{Tabelle}$: Nullhypothese wird akzeptiert
4. Interpretation des Ergebnisses Wenn die Nullhypothese verworfen wurde. Die Fehler-wahrscheinlichkeit für diese Entscheidung liegt (unter) der Irrtumswahrscheinlichkeit α. Wenn die Nullhypothese akzeptiert wurde. Die Fehler-wahrscheinlichkeit für diese Entscheidung (β-Fehler) ist nicht bekannt und über einen t-Test nicht ermittelbar.	**4. Interpretation des Ergebnisses** Wir gehen davon aus, dass sich die beiden Farbstoff-lösungen nicht in ihrem Gehalt unterscheiden. Die Unterschiede im Mittelwert der Bestimmung (Stich-proben) sind rein zufällig. Die exakte Wahrscheinlichkeit für eine fälschlicherweise Akzeptierung des Nullhypothese (β-Fehler) ist nicht bekannt.